INFORMATION AND CONTROL 26, 312–340 (1974)

Recovery of Randomly Sampled Signals by Simple Interpolators*

FREDERICK J. BEUTLER

Computer, Information and Control Engineering Program, The University of Michigan, Ann Arbor, Michigan 48104

The mean square error performance of simple polynomial interpolators is analyzed for wide-sense stationary signals subjected to randomly timed sampling represented by stationary point processes. This performance is expressed in dimensionsless parametric terms, with emphasis on asymptotic error behavior at high dimensionless sampling rates γ.

The form of the asymptotic error expression, and particularly its dependence on γ, is shown to vary according to the number of points utilized, together with the differentiability properties of the signal. One point extrapolation yields a mean square error varying with γ^{-2} if the signal is differentiable, and as γ^{-1} if the signal is not. Similarly, two-point (polygonal) interpolation error exhibits linearity in γ^{-4}, γ^{-3} or γ^{-2}, according as the signal is twice, exactly once, or nondifferentiable. Specific examples are offered to furnish insight into actual error magnitudes. It is shown, for instance, that introduction of jitter in the sampling sequence increases the error by only a negligible amount. Exponential decay of the sample values is compared with stepwise holding; little is gained for a nondifferentiable signal, while for a differentiable signal the error performance deteriorates from γ^{-2} to γ^{-1} at high sampling rates.

When more than two points are used in a polynomial fitting recovery scheme, specific computations or error become excessively difficult. However, it is proved that the asymptotic mean square error varies with γ^{-2n} when n points are utilized, and the signal is continuously differentiable at least n times.

Finally, we compare the mean square errors of one and two sample schemes as described above with those attained by causal (extrapolating) and noncausal (interpolating) Wiener-Kolmogorov optimal filters. We demonstrate nontrivial instances in which the Wiener-Kolmogorov mean square error varies as $\gamma^{-1/2}$, so that any of the simple recovery schemes considered exhibits superior performance at high sampling rates. This is explained by noting that the latter represent time-varying filters, whereas the Wiener-Kolmogorov filter is time-invariant.

* Research sponsored by the Air Force Office of Scientific Research, AFSC, USAF, under Grant No. AFOSR-70-1920B, and the National Science Foundation under Grant No. GK-20385.

I. Introduction

Practical recovery of a signal from its samples has concerned engineers for several decades, as Shannon's famous paper (1949) amply demonstrates. Originally, periodic sampling was routinely assumed. It was soon realized, however, that a stochastic model of the sampling sequence $\{t_n\}$ more realistically reflects errors (jitter) (Balakrishnan, 1962; Beutler and Leneman, 1968), lost samples (skipping), time-shared computer operation, and the like.

It is indeed tempting to approach reconstitution of the signal $x(t)$ through highly sophisticated optimization techniques. For instance, if $x(t)$ is wide-sense stationary and independent of (stationary) $\{t_n\}$, one can sometimes (see Brown, 1961; Leneman, 1966b) apply the Wiener-Kolmogorov technique to the sampled signal $y(t)$ by regarding it as a delta function train

$$y(t) = x(t)\, s(t) = \sum_{-\infty}^{\infty} x(t_n)\, \delta(t - t_n); \tag{1.1}$$

thus, one minimizes the mean square recovery error $\epsilon^2 = E\{[x(t) - \hat{x}(t)]^2\}$.[1]

In this work, we consider instead simple interpolators based on only a small number of samples, following Leneman (1966a), Leneman and Lewis (1966a and 1966b). It is our contention that such an interpolator is often preferable to a Wiener-Kolmogorov (or Kalman-Bucy) filter. Our preference follows from a number of considerations, namely:

1. The optimum is often difficult or impossible to compute.
2. Its implementation can be complicated.
3. Unless the statistics of $x(t)$ and $\{t_n\}$ are well known, a Wiener or Kalman-Bucy filter cannot be obtained. Even then, its effectiveness is seriously denigrated by noise and/or imperfect knowledge of the relevant statistics. On the other hand, simple interpolators are robust.
4. A simple interpolator depends neither on the remote past (as does the Wiener filter) nor on initial time varying computation and implementation (as does the Kalman-Bucy filter); it operates immediately and without start-up transient whenever the data becomes available.

Lastly—and most surprisingly—the simple interpolator often exhibits performance superior to a comparable Wiener filter. This apparent paradox (which the author first observed by a direct calculation) is readily explained.

[1] This optimization, as applied to generalized random processes, is rigorously treated in Gelfand and Vilenkin (1964) and Rozanov (1959).

A time invariant filter is perforce based only on the statistics of the sampling train $\{t_n\}$ only, whereas even a simple sample-and-hold interpolator represents a time-varying weighting function [acting linearly on $x(t)$] predicated on the actual realization of the sampling instants t_n [see Eq. (1.3)].

It is seen that simple interpolators are well worthy of attention. At the same time, our understanding of the performance of such interpolators (when used in connection with nonperiodic sampling) is quite limited. Answers do not now exist to any of the following:

1. For typical sampling schemes and random signals, what mean square error performance is achieved by various simple interpolators?

2. Can the error be expressed parametrically to show the extent of dependence on factors such as sampling rates, number of points employed, irregularity of sampling times, etc.?

3. What is the mean square interpolatory error compared with that attained by a Wiener-Kolmogorov filter (causal or noncausal)?

4. How is the error affected in form and magnitude by the characteristics of the signal, as expressed by its "smoothness" or degree of differentiability?

Although our results do not dispose of the above questions completely, we are able to make considerable progress with them. To do so, we limit ourselves to wide-sense stationary signals $x(t)$ stochastically independent of the sampling process $\{t_n\}$. Our $\{t_n\}$ is only subject to the weak hypothesis that $\{t_n\}$ is a stationary point process (hereafter abbreviated spp) (Beutler and Leneman, 1968, 1971), which imposes merely the requirement that any joint probability distribution of numbers of points in intervals be invariant under time axis translations. The class of spp thus embraces a wide variety of sampling schemes, including those most likely to appear in practice [see Beutler and Leneman (1968, 1971) for examples which include jitter, skips, Poisson, etc.]. Properties of spp are well understood (Beutler and Leneman, 1966a, 1966b), so that the analytical tools for investigating interpolation errors are at hand. Finally, the stationarity properties of $\{t_n\}$, in combination with those of $x(t)$, result in wide-sense stationary $\{x(t_n)\}$; then, unless the interpolation scheme is explicitly time dependent, the interpolated process $x(t)$ is wide-sense stationary also. Throughout our work, the measure of interpolator performance will be the mean square error $\epsilon^2 = E\{[x(t) - \hat{x}(t)]^2\}$. For convenience, we take $x(t)$ real, and to standardize our results, we shall always suppose $E\{[x(t)]^2\} = 1$.

We begin with the simplest possible interpolator, the sample-and-hold scheme shown in Fig. 1, and described by the output

$$\hat{x}(t) = x(t_n) \qquad \text{for} \quad t_n \leqslant t < t_{n+1}\,. \tag{1.2}$$

Each $x(t_n)$ is thus held until the arrival of the next sample at time t_{n+1}, at which instant the current $\hat{x}(t)$ is "dumped" and replaced by a new value $x(t_{n+1})$. While the sample-and-hold interpolator corresponds to the time-varying weighting function[2]

$$h(t, u) = \sum_{-\infty}^{\infty} [U(u - t_n -) - U(u - t_n +)][U(t - t_n) - U(t - t_{n+1})], \tag{1.3}$$

it is actually easy to mechanize as a hardware device, and its performance can readily be calculated without reference to (1.3).

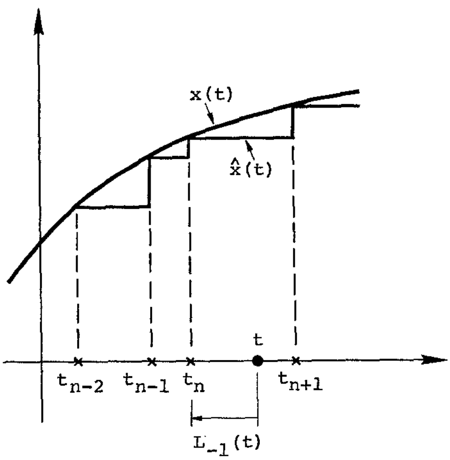

FIG. 1. Sample-and-hold interpolation.

A variation of the above method, designated exponential interpolation, is shown in Fig. 2; it differs from sample-and-hold only in the exponential weighting of the sample value used. In other words,

$$\hat{x}(t) = e^{-b(t-t_n)}x(t_n) \qquad \textit{for} \quad t_n \leqslant t < t_{n+1}\,, \tag{1.4}$$

where $b > 0$ is chosen as desired. On strictly intuitive grounds, it could be

[2] Here $U(\cdot)$ is the unit step function.

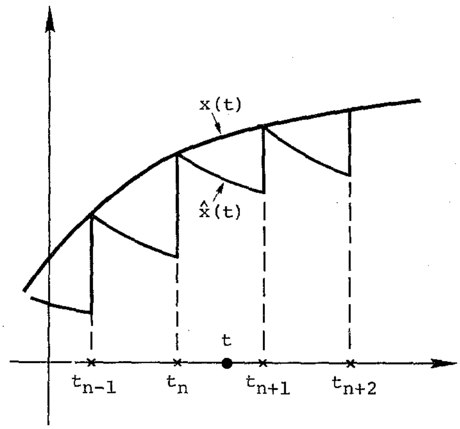

FIG. 2. Exponential decay interpolation.

argued that the exponential sampling of (1.4) is preferable to (1.2), since exponential sampling is superior at low sampling rates. In fact, if the autocorrelation of $R_x(\tau) \to 0$ as $\tau \to \infty$ [always true if $x(t)$ has an absolutely continuous spectral distribution function], we have $\epsilon^2 \to 2$ as the sampling rate tends toward zero with sample-and-hold recovery, whereas $\epsilon^2 \to 1$ for exponential interpolation.

At high sampling rates, however, the relative ranking of these two interpolation schemes is reversed for all differentiable signals. This becomes clear once exponential interpolation is analyzed in conjunction with sample-and-hold interpolation by calculating errors respective to (1.4) with arbitrary $b \geqslant 0$, that is, by regarding sample-and-hold as a special or limiting case of exponential interpolation.

On each interval $[t_n, t_{n+1})$, sample-and-hold recovery is recognized as polynomial interpolation [see Davis (1963, p. 24)] with $n = 1$. The same type of interpolation, but with $n = 2$, connects successive samples with straightline segments as shown in Fig. 3. We refer to the interpolatory method of Fig. 3 as polygonal interpolation. When the sampling times are random, computation of mean square errors becomes quite difficult; this becomes evident when we see that recurrence times (i.e., time intervals from t to the nearest sampling instant) are present in both numerator and denominator of a quotient, and that joint statistics of forward and backward recurrence times are required. Nonetheless, we have been able to determine the mean square error for some special cases, and to formulate parametric relations on error behavior. The latter lend themselves naturally to comparisons with noncausal Wiener-

Kolmogorov filtering, since the usual polygonal interpolator is itself noncausal. It will be observed that the simple interpolator exhibits better performance at high sampling rates, even though its delay requirements are small, in contrast to the infinite delay postulated for the noncausal Wiener-Kolmogorov filter.

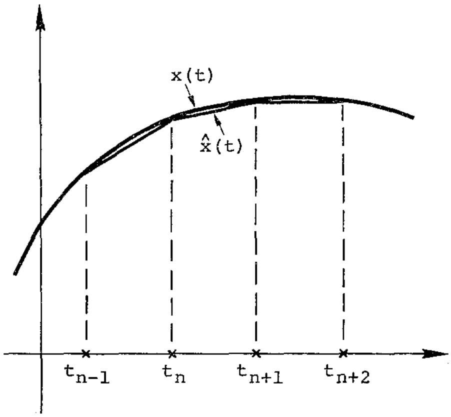

FIG. 3. Polygonal interpolation.

We also consider both causal and noncausal polynomial interpolators of arbitrary degree at high sampling rates. Exact mean square error expressions then become insuperably difficult to calculate, and in any case, lead to such complicated expressions that insight is inevitably lost. At high sampling rates, however, simple parametric relationships are obtained, and easy mean square error bounds established.

It should be noted that the emphasis on high sampling rate results is not a serious restriction. Indeed, any meaningful application of sampled signal recovery requires that the error be small relative to the signal magnitude, and this in turn demands high sampling rates. The notion of sampling rate is made more precise in the next section, in which some of the other implication of asymptotic error analysis at high sampling rates is also discussed.

II. NOTATION AND BACKGROUND

Since the material of this paper leans heavily on properties of spp, it is convenient to summarize the notation and background to be used; the interested reader will find further results and the applicable derivations in Beutler and Leneman (1966a, 1966b). Fundamental is the notion of recurrence time $L_k(t)$, which is the time interval between t and the kth sampling point occurring thereafter. If k is a negative integer, $L_k(t)$ denotes the interval length between the kth point preceding t and t itself. The distribution function of $L_k(t)$ does not depend on t (by virtue of the stationarity of $\{t_n\}$), and is known to be absolutely continuous; hence, a statistical description of $L_k(t)$ is furnished by the probability density function $g_k(\cdot)$, which is zero for negative argument. A particularly prominent role is played by g_1, which turns out to be the density of the distance to the nearest sampling point in either direction. Moreover, g_1 is a monotone function for which $g_1(0+) = \beta$, where β is defined to be the mean number of sampling points per unit time.

It is sometimes convenient as well as natural to operate in the frequency domain, for which purpose one introduces the transform

$$g_k^*(s) = \int_0^\infty g_k(u)\, e^{-su}\, du = E\{\exp[-sL_k(t)]\}. \tag{2.1}$$

It can be shown that g_k^* is related to the conditional distribution F_k for $L_k(t)$, given that a sampling point occurs at t; if f_k^* is the Laplace-Stieltjes transform of F_k, we have [see Beutler and Leneman (1968b)]

$$g_k^*(s) = \beta s^{-1}[f_{k-1}^*(s) - f_k^*(s)]. \tag{2.2}$$

If the intervals are independent and identically distributed, F_k becomes the (unconditioned) distribution of k successive intervals.

Of greatest interest are the recovery capabilities of interpolation schemes at relatively high sampling rates. To justify this claim, it would suffice to remark that the mean square error (i.e., error power) should be small compared to the signal power in any practical situation, and that this is true only at high sampling rates. However, there are additional reasons for concentrating on high sampling rates. For one thing, the mean square error expressions tend to be quite complex, so that comparison among sampling schemes, signals, and interpolators is awkward. The asymptotic (high sampling rate) formulas are much simpler, so that a direct parametric understanding of factors affecting error becomes possible. Finally, one can often find universal error formulations related to moments of $x(t)$ and $\{t_n\}$, independently of the form of the autocorrelation or other statistics.

Insofar as possible, we shall present our results in nondimensionally normalized form. For instance, our notion of sampling rate is the nondimensional ratio $\gamma = \beta/W$, in which β is the mean number of samples per unit time (as previously defined), and W is some consistent bandwidth measure for $x(t)$. One possible choice of W is

$$W = \inf_{A} \left\{ \frac{1}{2\pi} \int_{-A}^{A} S_x(\omega)\, d\omega = \frac{1}{2} \right\}, \tag{2.3}$$

S_x being the spectral density of $x(t)$.[3] The mean square error will then be given in terms of the (normalized) sampling rate γ, together with some statistical parameters of the unit bandwidth process $\tilde{x}(Wt) = x(t)$ and the unit rate sampling sequence $\{\tilde{t}_n\}$ defined by $\tilde{t}_n = \beta t_n$. In this fashion, it is possible to isolate the effect of sampling rate on the error, as distinct from variations in $x(t)$ and $\{t_n\}$ unrelated to frequency content. If the mth moment of the nth derivative [of $x(t)$ in quadratic mean] exists, the separation is obtained from relations such as

$$E\{[x^{(n)}(t)]^m\} = W^{nm} E\{[\tilde{x}^{(n)}(0)]^m\} \tag{2.4}$$

and for the m th moments of the recurrence time $L_k(t)$

$$E\{[L_k(t)]^m\} = \beta^{-m} E\{[\tilde{L}_k(0)]^m\}. \tag{2.5}$$

Elementary calculations also yield $g_k(u) = \beta \tilde{g}_k(\beta u)$, $g_k^*(s) = \tilde{g}_k^*(s/\beta)$, $S_x(\omega) = W^{-1}\tilde{S}_x(W^{-1}\omega)$ and $R_x(\tau) = R_{\tilde{x}}(W\tau)$; here we have consistently used the tilde to refer to statistics of the normalized processes, and have denoted correlations by R with appropriate subscript.

Although some of our results hold for any spp satisfying $E\{[L_1(0)]^2\} < \infty$, it is frequently desirable and necessary to turn to explicit forms, chosen on the basis of analytic tractability as well as practical applicability. One of these is the so-called skip-jitter sampling process, which can be viewed as idealized periodic sampling (period T) modified by equipment and environmental conditions. One supposes that points are deleted independently, each with deletion probability q. In addition, the remaining sample points are randomly perturbed; jitter displacements u_k are mutually independent, and may be statistically described by the generating function[4]

$$C(s) = E[\exp(-s u_k)]. \tag{2.6}$$

[3] We assume throughout for convenience only that the spectral distribution function of $x(t)$ is absolutely continuous.

[4] The jitter cannot be so large that indices of sampling points are permuted. It follows that the u_k must be distributed over an interval whose length cannot exceed T, and which (by the stationarity of $\{t_n\}$) can be taken as $[0, T)$.

With the above notation, we obtain for the skip-jitter process,

$$f_n^*(s) = C(s)\,C(-s)[(1-q)\,e^{-sT}]^n/[1-qe^{-sT}]^n, \tag{2.7}$$

as shown in Beutler and Leneman (1966b), Eq. (7.20). Of course, the skip-jitter process can be specialized if desired by setting either $q = 0$ or $C(s) = 1$. For comparison, it is of interest to consider the "most random" [i.e., largest entropy—see McFadden (1965)], spp, namely, the well-known Poisson point process. For this spp

$$g_1(u) = \beta e^{-\beta u} \qquad \text{and} \qquad g_n^*(s) = f_n^*(s) = \left[\frac{\beta}{s+\beta}\right]^n, \tag{2.8}$$

which corresponds to the classical formulas of renewal theory.

One process $x(t)$ whose interpolation is analyzed in considerable detail is the wide sense Markov process, whose correlation is necessarily of the form [see Beutler (1963), Theorem 3]

$$R_x(\tau) = \exp(-a\mid\tau\mid). \tag{2.9}$$

There are several reasons why this stochastic process should be of special interest to us. Aside from its mathematical simplicity, we note that the correlation is not differentiable at the origin, so that $x(t)$ is nondifferentiable in quadratic mean. Thus, the wide sense Markov process is suitable to the study of interpolation for nondifferentiable processes. Furthermore, it is known that the minimum mean square causal estimator for the process is (Beutler, 1961)

$$\hat{x}(t) = x(t - L_{-1}(t))\exp[-aL_{-1}(t)], \tag{2.10}$$

with resulting mean square error

$$\epsilon^2 = 1 - g_1^*(2a) \tag{2.11}$$

for any spp sampling process whatsoever. Lastly, the interpolation error for any $x(t)$ with rational spectral density having distinct poles is readily found once the error corresponding to (2.9) is known. This follows because the mean square error is linear in R_x, so that a solution for (2.9) extends immediately to $x(t)$ possessing correlations of the form

$$R_x(\tau) = \sum_k A_k[\exp(-a_k\mid\tau\mid)]. \tag{2.12}$$

III. Exponential and Sample-and-Hold Interpolation

In this section, we obtain and compare the mean square errors incurred when the recovery is based on the single sample value which immediately precedes the present time t. We consider in particular interpolators whose output is

$$\hat{x}(t) = x(t - L_{-1}(t)) \exp[-bL_{-1}(t)] \tag{3.1}$$

for some $b \geqslant 0$, as in (1.2) and (1.4). Because the statistics of $\{t_n\}$ appear only in a simple form within (3.1), the mean square error corresponding to the $\hat{x}(t)$ of (3.1) is relatively easy to evaluate.

For interpolators or estimators of general type, the mean square error can always be written as

$$\epsilon^2 = 1 - 2E[x(t)\,\hat{x}(t)] + E\{[\hat{x}(t)]^2\}, \tag{3.2}$$

where $x(t)$ and $\hat{x}(t)$ are assumed to be of finite mean square, with $x(t)$ normalized so that $E\{[x(t)]^2\} = 1$. Since $x(t)$ and $\{t_n\}$ are supposed statistically independent, expectations involving $x(t)$ can be calculated when desired by applying successively these expectations in the more convenient order. We thus have

$$\begin{aligned} E[x(t)\,\hat{x}(t)] &= E\{R_x[L_{-1}(t)] \exp[-bL_{-1}(t)]\} \\ &= \int_0^\infty R_x(u)\, e^{-bu} g_1(u)\, du \end{aligned} \tag{3.3}$$

and (by an even simpler computation)

$$\begin{aligned} E\{[\hat{x}(t)]^2\} &= E\{[x(t)]^2\}\, E\{\exp[-2bL_{-1}(t)] \\ &= \int_0^\infty e^{-2bu} g_1(u)\, du = g_1{}^*(2b). \end{aligned} \tag{3.4}$$

From the three foregoing equations, it is seen that the mean square error pertinent to the exponential interpolator is

$$\varepsilon^2 = \int_0^\infty [1 - 2R_x(u)\, e^{-bu} + e^{-2bu}]\, g_1(u)\, du \tag{3.5}$$

for arbitrary $x(t)$ and $\{t_n\}$. The expression (3.5) is especially simple for a wide sense Markov signal.

EXAMPLE 3.1 (wide-sense Markov signal–arbitrary sampling statistics). On substituting in (3.5), we find that

$$\epsilon^2 = 1 - 2g_1^*(a + b) + g_1^*(2b). \tag{3.6}$$

If W is defined as in (2.3), we have $W = a$. We also let $\alpha = b/a$, so that the error takes on the dimensionless form

$$\epsilon^2 = 1 - 2\tilde{g}_1^*(\gamma^{-1} + \alpha\gamma^{-1}) + \tilde{g}_1^*(2\alpha\gamma^{-1}). \tag{3.7}$$

At high sampling rates, γ becomes large, and since (see Section II) $\alpha = 1$ minimizes the mean square error, we may suppose α (the normalized exponential decay) not large. Then we may take advantage of the small arguments of $\tilde{g}_1^*$ to derive a high sampling rate approximation to (3.7). In fact, the assumed existence of the second moment of $L_1(t)$ implies [see Loeve (1955, p. 199)] that

$$g_1^*(z) = 1 - zE[L_1(0)] + \tfrac{1}{2}z^2E\{[L_1(0)]^2\} + o(z^2) \tag{3.8}$$

where $o(z^2)/z^2 \to 0$ as $z \to 0$ by definition. With the substitution (3.8), (3.7) becomes

$$\epsilon^2 = 2\gamma^{-1}E[\tilde{L}_1(0)] + \gamma^{-2}[\alpha^2 - 2\alpha - 1]\, E\{[\tilde{L}_1(0)]^2\} + o(\gamma^{-2}). \tag{3.9}$$

The significance of (3.9) is discussed below, when a comparable expression for differentiable signals is available. To this end, we first derive an alternative frequency domain version of (3.5). One of the terms needed is already given by (3.4). The other follows from (3.3) by writing R_x as an inverse Fourier transform of S_x, and then (since g_1 and S_x are both integrable) applying Fubini's theorem. The final result, stated in dimensionless parameters, is

$$\epsilon^2 = 1 + \tilde{g}_1^*(2\alpha\gamma^{-1}) - \frac{1}{\pi}\int_{-\infty}^{\infty} S_{\tilde{x}}(\omega)\, \tilde{g}_1^*(\alpha\gamma^{-1} - i\omega\gamma^{-1})\, d\omega. \tag{3.10}$$

We use (3.10) in

EXAMPLE 3.2 (differentiable signal—high sampling rates). If $\tilde{g}_1^*$ in (3.10) is put in series form (3.8), there follows

$$\epsilon^2 = \gamma^{-2}E\{[\tilde{L}_1(0)]^2\}(\alpha^2 + E\{[\dot{\tilde{x}}(0)]^2\}) + o(\gamma^{-2}). \tag{3.11}$$

We conclude that the mean square error always varies as the inverse square of the sampling rate when the signal is differentiable; this is clearly more desirable than the variation with γ^{-1} encountered for the wide-sense

Markov signal. We further conjecture—but have been unable to prove—that the error behaves as γ^{-1} for any nondifferentiable signal.

Exponential interpolation is undesirable in conjunction with a differentiable signal $x(t)$; as (3.11) indicates, the mean square error increases directly with the decay rate squared. On the other hand, the error (3.9) applicable to the wide sense Markov signal is quite insensitive to the decay rate α even at moderate sampling rates γ^{-1}. Both error behavior with γ and effect of decay are readily explained by the manner in which R_x drops off near the origin. If $x(t)$ is differentiable, $\dot{R}_x(0) = 0$, and so $\hat{x}(t)$ is more highly correlated with $x(t)$ than for the nondifferentiable (wide-sense Markov) signal. This leads to smaller interpolation errors for differentiable signals, and also suggests that decay in $\hat{x}(t)$ emphasizes its deviation from $x(t)$ whenever R_x departs only a little from unity during a sampling interval.

Our next object of investigation is the influence of jitter on the mean square error. A limited approach is through (3.9) or (3.11), in which one observes the change in $E[\tilde{L}_1(0)]$ or $E\{[\tilde{L}_1(0)]^2\}$, respectively, as the sampling point locations are modified by independent identically distributed perturbations.

Example 3.3 (wide-sense Markov signal—jittered periodic sampling). The interval lengths in this instance are $T + u_n - u_{n-1}$, where the u_k are the jitter variables. Then if $Q(\cdot)$ is the (common) distribution function of $u_n - u_{n-1}$, we have $g_1(u) = T^{-1}[1 - Q(u - T)]$ by the relation of F_1 to g_1 [see Beutler and Leneman (1966a), Eq. (3.4.1)]. One then computes

$$E\{[\tilde{L}_1(0)]^m\} = \int_0^\infty u^m \tilde{g}_1(u)\, du.$$

We shall assume in particular that each u_k is uniformly distributed over the interval $[0, \nu T)$, whence it follows that

$$E[\tilde{L}_1(0)] = \frac{1}{2} + \frac{\nu^2}{12}; \tag{3.12}$$

the mean square error (3.9) is therefore $\epsilon^2 = \gamma^{-1}[1 + (\nu^2/6)] + o(\gamma^{-2})$. Although jitter augments the error, the increase is negligibly small; if each sampling point is allowed to deviate by as much as a quarter of the sampling interval (i.e., $\nu = \frac{1}{4}$), there is an increase of but 1% in the error. The same reasoning is pertinent to differentiable signals, for which we have:

Example 3.4 (differentiable signal—jittered periodic sampling). We proceed as in Example 3.3, and obtain $E\{[\tilde{L}_1(0)]^2\} = \frac{1}{3} + (\nu^2/6)$. The latter yields a high sampling rate approximation

$$\epsilon^2 = \frac{1}{3}\gamma^{-2}\left(1 + \frac{\nu^2}{2}\right)(\alpha^2 + E\{[\dot{x}(0)]^2\}) + o(\gamma^{-2}). \tag{3.13}$$

If the sampling points are perturbed in a uniform distribution extending (as in Example 3.3) over a quarter of the sampling interval, the error increase is 3 %—still a small amount, especially in view of the relatively large sampling time deviation. The small influence of jitter can be explained intuitively by observing that $L_{-1}(t)$ is unchanged on the average, but since there is a symmetrical probability of smaller and larger occurrence times before the most recent sample, the mean square weighting generates a second order error effect.

Jitter may be imposed on sampling processes other than the periodic. It is seen from (3.7) and (3.10) that the change in $\tilde{g}_1{}^*$ resulting from the jitter translates into its influence on the mean square error. Consequently, general approximation formulas can be derived if the variation in $g_1{}^*$ can be related to the jitter. Let us call f_{01}^* the interval characteristic function [see Eq. (2.2)] prior to the imposition of jitter, and take $C(\cdot)$ the jitter characteristic function as defined by (2.6). Then

$$f_1{}^*(s) = C(s)\,C(-s)\,f_{01}^*(s) \tag{3.14}$$

and from the connection (2.2) between $f_1{}^*$ and $g_1{}^*$

$$\Delta g_1{}^*(s) = g_1{}^*(s) - g_{01}^*(s) = \frac{1 - C(s)\,C(-s)}{s}\,[sg_{01}^*(s) - 1]. \tag{3.15}$$

In nondimensional form, (3.15) can be expanded in the series,

$$\Delta \tilde{g}_1{}^*(s) = -s\sigma^2 + s^2\sigma^2 + o(s^2), \tag{3.16}$$

in which σ^2 is the variance of the normalized jitter variable βu_k.

EXAMPLE 3.5 (wide-sense Markov signal—arbitrary jitter statistics). The mean square error increase consequent to the introduction of jitter is determined for moderate or larger sampling rates by substituting (3.16) into the general mean square error formula (3.7). Upon performing the indicated operations, one sees that

$$\Delta\epsilon^2 = 2\sigma^2\gamma^{-1} + 2\sigma^2\gamma^{-2}(\alpha^2 - 2\alpha - 1) + o(\gamma^{-2}). \tag{3.17}$$

Evidently, jitter enlarges the error by the ratio $\sigma^2/E[\tilde{L}_1(0)]$ for high sampling rates. Although we have been unable to find an *a priori* bound for this ratio, the requirement that the jitters not permute sampling points suggests that 1/4 is not exceeded. As we have seen in Examples 3.3 and 3.4, even large jitters do not come close to this value under plausible conditions.

The techniques of Example 3.5 are equally suitable to a determination of error increase for differentiable signals when jitter is imposed on the sampling sequence.

EXAMPLE 3.6 (differentiable signal—arbitrary jitter statistics). We put (3.16) into (3.10), observing that some terms cancel, while others assume the form $\int \omega S_x(\omega)\, d\omega = 0$. There remains

$$\Delta\epsilon^2 = 2\sigma^2\gamma^{-2}(\alpha^2 + E\{[\mathring{x}(0)]^2\}) + o(\gamma^{-2}). \tag{3.18}$$

In this instance, the mean square error augmentation ratio is $2\sigma^2/E\{[\tilde{L}_1(0)]^2\}$, which is again likely to be small in practical applications.

Skipping of samples is yet another phenomenon which may be encountered in practice as a result of sampler malfunction, jamming, fading or other environmental causes. Like jitter, skipping exacts a performance penalty intuitively connected with irregularity of sampling intervals. For simplicity, we consider only periodic sampling subjected to independent deletions, each sample being omitted with probability q. The $f_1{}^*$ for this case is then given by (2.7) with $C(s) = 1$, and

$$E\{[L_1(0)]^m\} = \frac{(-1)^{m+1}\beta}{m+1} \frac{d^{m+1}}{ds^{m+1}} f_1{}^*(s) \Big|_{s=0} \tag{3.19}$$

where $\beta = (1 - q)/T$. We are therefore led to the high sampling rates results of

EXAMPLE 3.7 (periodic sampling with independent random skips). Because $E[L_1(0)] = \frac{1}{2}(1 + q)\,\beta^{-1}$, substitution in (3.9) for the wide-sense Markov signal yields

$$\epsilon^2 = \gamma^{-1}(1 + q) + o(\gamma^{-1}). \tag{3.20}$$

The procedure is entirely analogous for differentiable signals; taking $m = 2$ in (3.19) and putting $E\{[\tilde{L}_1(0)]^2\}$ into (3.11) gives

$$\epsilon^2 = \tfrac{1}{3}\gamma^{-2}(1 + 4q + q^2)(\alpha^2 + E\{[\mathring{x}(0)]^2\}) + o(\gamma^{-2}). \tag{3.21}$$

The error increase engendered by the skipping is connected with the consequent irregularity of the sampling, and is expressed by the factors $(1 + q)$ and $(1 + 4q + q^2)$ in (3.20) and (3.21), respectively. Of course, if T is held constant in face of increasing skipping probability q, the sampling rate γ decreases, thus leading to a larger error also. It is worth noting that the sampling tends to a Poisson process if we let $q \to 1$ while taking $T \to 0$ such that the ratio $(1 - q)/T$ remains undisturbed; we could use this fact to calculate mean square errors attendant to Poisson sampling, but find it as easy to proceed directly.

We next compare the mean square error of exponential interpolation with that achieved by causal Wiener-Kolmogorov filtering. Unfortunately, the comparison is severely limited by difficulties in the factorization which is a crucial aspect of the Wiener procedure. Nonetheless, our results strongly suggest the conjecture that the exponential interpolator gains in relative performance as the sampling sequence becomes more random.

Our discussion is of necessity confined to wide-sense Markov signals. For these, the exponential interpolator (with $\alpha = 1$) is optimal (among all linear causal filters) with respect to *any* sampling sequence (Beutler, 1961). When the sampling sequence is periodic, the exponential interpolator and Wiener filter coincide. But if random skipping (see preceding example) is introduced, the corresponding Wiener filter possesses infinite memory (Leneman, 1966b);[5] since its output must then differ from the optimum (exponential) interpolator (see Beutler, 1961), the causal Wiener filter must generate a larger mean square error.

An explicit comparison is feasible for Poisson sampling, which can be regarded as "most random" among all sampling schemes when judged on the basis of entropy (McFadden, 1965).

EXAMPLE 3.8 (wide-sense Markov signal—Poisson sampling). The exact mean square error formula is immediate from (2.8) and (3.7), viz,

$$\epsilon^2 = 2\left\{\frac{1 + \alpha\gamma^{-1}}{1 + 2\alpha\gamma^{-1}} - \frac{1}{1 + \alpha\gamma^{-1} + \gamma^{-1}}\right\}. \tag{3.22}$$

The asymptotic error form is computed either from (3.22), or by using $E[\tilde{L}_1(0)] = 1$ in (3.9), whence

$$\epsilon^2 = 2\gamma^{-1} + o(\gamma^{-1}). \tag{3.23}$$

[5] In this instance, the Wiener filter and its mean square error depend implicitly on q, T and a through the solution of a quadratic equation. We have therefore been unable to obtain a high sampling rate approximation to the mean square error.

This represents a definite improvement over Wiener filtering, for which (Leneman, 1966b, Eq. 2.43) we have the asymptotic error relation

$$\epsilon^2 = \sqrt{2}\,\gamma^{-1/2} + o(\gamma^{-1}). \tag{3.24}$$

Further quantitative comparisons between Wiener filters and simple interpolators appear in the next section. There, the fair comparison is with noncausal time-invariant mean square optimal filters. Since the latter are calculable without the need for spectral factorization, we will be able to obtain a greater number of results for reference.

IV. Polygonal Interpolation

The interpolatory methods presented thus far utilize only the last sample preceding time t to produce $\hat{x}(t)$. Use of a greater number of samples in linear combination should decrease the interpolation error; this is expected in analogy with the nonrandom interpolation of functions [see, for instance, Davis (1963), p. 56 ff.]. There we recognize sample-and-hold interpolation as zeroth order polynomial interpolation, while polygonal interpolation is identified with interpolation by polynomials of order one. It follows that the sample-and-hold error should depend inversely on the interval length, while the error of polygonal interpolation is inverse to the square of this length. Moreover, the error for the latter becomes much smaller (by a constant multiplicative factor) when one sample is to each side of time t than when both samples lie in the past.

Motivated by the above considerations, we now analyze polygonal interpolation; it is of particular interest to note whether the behavior of the mean square error is consistent with the arguments of the preceding paragraph. To this end, consider the interpolation scheme illustrated by Fig. 3. We see that $\hat{x}(t)$ lies on the straight line joining the samples $x(t_n)$ and $x(t_{n+1})$, the index n being such that $t_n \leqslant t < t_{n+1}$. Evidently, this interpolation method is noncausal, although the delay is likely to be small at high sampling rates.

An analytical expression for $\hat{x}(t)$ can be developed in terms of the forward recurrence time $L_1(t)$ and backward recurrence time $L_{-1}(t)$; the contributions of $x(t_n)$ and $x(t_{n+1})$ are in linear combination that depends on these recurrence times. More specifically, we have for $\hat{x}(t)$

$$\hat{x}(t) = \frac{L_{-1}x(t + L_1) + L_1 x(t - L_{-1})}{L_1 + L_{-1}}, \tag{4.1}$$

in which we have suppressed the arguments of $L_1(t)$ and $L_{-1}(t)$.

Calculation of the mean square error ϵ^2 from (3.2) requires a knowledge of $E\{[\hat{x}(t)]^2$ and $E[x(t)\,\hat{x}(t)]$, where $\hat{x}(t)$ is given by (4.1) above. We cannot in general hope to obtain these expectations without the joint probability distribution of L_1 and L_{-1}, since these appear in nonlinear combination in the expression for the square of the error. Even when the appropriate integrands are available, the $(L_1 + L_{-1})$ term in the denominator of (4.1) usually leads to integrals that cannot be evaluated in closed form. In view of these problems, few direct results are explicitly obtainable.

It is advantageous to use the symmetry of R_x to write

$$E[x(t)\,\hat{x}(t)] = E\left\{\frac{(L_{-1})\,R_x(L_1) + (L_1)\,R_x(L_{-1})}{L_1 + L_{-1}}\right\} \tag{4.2}$$

and to put $E\{[\hat{x}(t)]^2\}$ in the form

$$E\{[\hat{x}(t)]^2\} = E\left\{\frac{(L_{-1})^2 + (L_1)^2 + 2(L_1L_{-1})\,R_x(L_1 + L_{-1})}{(L_1 + L_{-1})^2}\right\}. \tag{4.3}$$

We can then obtain a fairly general error formula, namely

Example 4.1 (twice differentiable signal—high sampling rates). Let us call $G(\cdot, \cdot)$ the joint distribution function of L_1 and L_{-1}, and assume the existence of the joint second moments of the recurrence times, viz,

$$E\{[L_1(t)\,L_{-1}(t)]^2\} = \iint_0^\infty u^2v^2\,dG(u, v). \tag{4.4}$$

The mean square error, as furnished by (3.2), now appears as a double integral with respect to G, with $E[x(t)\,\hat{x}(t)]$ and $E\{[\hat{x}(t)]^2\}$ given by (4.2) and (4.3), respectively. Because $x(t)$ is supposed twice differentiable, R_x may be expanded as in Loeve (1955, p. 199). If the integral is written in dimensionless form using the sampling rate γ, we find that

$$\epsilon^2 = \tfrac{1}{4}\gamma^{-4}E\{[\ddot{x}(t)]^2\}\,E\{[\tilde{L}_1(t)\,\tilde{L}_{-1}(t)]^2\} + o(\gamma^{-4}). \tag{4.5}$$

If $G(\cdot, \cdot)$ is too difficult to find, one may modify (4.5) by utilizing the upper bound

$$E\{[\tilde{L}_1(t)\,\tilde{L}_{-1}(t)]^2\} \leqslant E\{[\tilde{L}_1(t)]^4\}, \tag{4.6}$$

the latter follows from the Schwarz inequality and the fact that L_1 and L_{-1} both have the same probability distribution [see Beutler and Leneman (1966a)].

It is noteworthy that the mean square error varies with the sampling rates according to γ^{-4}. This behavior, together with that of once differentiable signals under sample-and-hold interpolation, is entirely consistent with the classical results on polynomial interpolation [see again Davis (1963)]. While it would be desirable to compare the mean square error expression (4.5) with the corresponding error for Wiener-Kolmogorov filtering, nontrivial computations on the latter appear to pose insuperable difficulties.

As we have observed earlier, the polygonal interpolator of Example 4.1 is noncausal. However, this fact is irrelevant to the fashion in which the mean square error varies with the sampling rate γ. To further demonstrate this fact, and to observe the influence of causality, we borrow a result from the next section for:

EXAMPLE 4.2 (twice differentiable signal—causal polygonal extrapolator at high sampling rates). We suppose here that $\hat{x}(t)$ is the linear extrapolation of the two sample values immediately preceding t, i.e.,

$$\hat{x}(t) = x(t - L_{-1}) + \left\{\frac{x(t - L_{-1}) - x(t - L_{-2})}{L_{-1} - L_{-2}}\right\}(t - L_{-1}). \tag{4.7}$$

Then it may be shown that

$$\epsilon^2 = \tfrac{1}{4}\gamma^{-4}E\{[\ddot{x}(0)]^2\}\,E\{[\tilde{L}_{-1}(t)\,\tilde{L}_{-2}(t)]^2\} + o(\gamma^{-4}); \tag{4.8}$$

it is clear that the only difference between the errors of the noncausal and causal interpolators lies in the mean square distance between t and the sampling instants.

For purposes of comparison, it is of interest to find also the form and error magnitude associated with optimum two point interpolation. That is, we take

$$\hat{x}(t) = a(t)\,x(t - L_{-1}) + b(t)\,x(t + L_{-1}), \tag{4.9}$$

choosing a and b to minimize the mean square interpolation error. Applying the orthogonality principle, one easily shows

$$a(t) = \frac{R_x(L_{-1}) - R_x(L_1)\,R_x(L_1 + L_{-1})}{1 - [R_x(L_1 + L_{-1})]^2}, \tag{4.10}$$

and that (by symmetry) b is precisely like a, but with L_1 and L_{-1} interchanged. The mean square error for this (optimal) a and b is calculated asymptotically at high sampling rates by substituting (4.9) into $\epsilon^2 = 1 - E[x(t)\,\hat{x}(t)]$ with optimal values of a and b entered therein. Although the formal computation is routine, as are the simpler ones leading to mean square errors (4.5) and (4.8), rigorous justification of the results prove more complicated in all these cases. The problems arise relative to the truncated Maclaurin expansion of R_x, since the latter is expressed in series form prior to taking the expectation with respect to $\{t_n\}$; it must then be shown that the remainder term leads only to a contribution of $o(\gamma^{-4})$ in the final expression. The denominator term of (4.10) necessitates particularly delicate analysis, since the manner in which it tends toward zero with $L_1 + L_{-1}$ is crucial. Under mild assumptions (i.e., the existence of the sixth moment of L_1), one shows the total integrand with respect to the probability distribution $G(\cdot, \cdot)$ well behaved, and indeed admitting an asymptotic expansion whose leading term has γ^{-4} as its multiplicand. We shall omit the rigorous verification of this assertion, and only note that the final result is

$$\epsilon^2 = \tfrac{1}{4}\gamma^{-4}(E\{[\ddot{x}(t)]^2\} - (E\{[\dot{x}(t)]^2\})^2)\, E\{[\tilde{L}_1(t)\tilde{L}_{-1}(t)]^2\} + o(\gamma^{-4}). \quad (4.11)$$

Thus, there is some reduction in mean square error, but no fundamental difference in error behavior with sampling rate γ, or with the sampling scheme [reflected only by the last term on the right of Eq. (4.11)]. Since a and b must be individually computed for every t, it seems questionable whether the additional complexity of the two point optimization is warranted.

To secure a better feeling for the comparative performance of the two point interpolators discussed above, we shall compute their respective mean square errors under similar conditions. To this end, let us assume Poisson sampling of a band-limited white signal.

EXAMPLE 4.3 (band-limited white signal with Poisson sampling—comparison of two-point reconstruction of signal). It is only necessary to substitute in the expressions already derived. Straightforward calculations yield

$$\begin{gathered} E\{[\ddot{x}(t)]^2\} = \frac{16}{5}, \qquad (E\{[\dot{x}(t)]^2\}) = \frac{4}{3},^{6} \\ E\{[\tilde{L}_1(t)\tilde{L}_{-1}(t)]^2\} = 4 \qquad \text{and} \qquad E\{[\tilde{L}_{-1}(t)\tilde{L}_{-2}(t)]^2\} = 40. \end{gathered} \quad (4.12)$$

For the (non-causal) polygonal interpolator, we then have

$$\epsilon^2 = \frac{16}{5}\gamma^{-4} \quad (4.13)$$

[6] In this example, our notion of bandwidth is furnished by the half power frequency, i.e. by (2.3). As usual, we suppose $x(t)$ normalized to make $E\{[x(t)]^2\} = 1$.

whereas the causal polygonal extrapolator yields 10 times this value or

$$\epsilon^2 = 32\gamma^{-4}. \tag{4.14}$$

As is expected, the smallest mean square error is associated with the optimal two-point interpolator; it is

$$\epsilon^2 = \frac{64}{45}\gamma^{-4}. \tag{4.15}$$

We again call attention to the similarity of the error forms, and that the dependence on sampling rate is the same for all. Indeed, differences in signal or sampling characteristics, as well as the choice of two-point reconstruction method, is reflected only in a fixed numerical multiplier.

The superiority of the simple two-point interpolator over a Wiener-Kolmogorov noncausal filter becomes evident when we calculate the mean-square error for the latter under comparable conditions. We again assume Poisson sampling of band-limited white signal. The technique (standard in any case) suggested in Leneman (1966b) is applied to the spectral densities obtained from Beutler and Leneman (1968), Eq. (3.7) and (4.6). The mean square optimal time-invariant interpolator then yields mean square error

$$\epsilon^2 = \frac{2}{\pi}\gamma^{-1} + o(\gamma^{-2}),^7 \tag{4.16}$$

which compares unfavorably with the variation with γ^{-4} of any of the two point interpolators mentioned in Example 4.3. Nor is the behavior of the Wiener-Kolmogorov interpolator merely a special case relevant only to Poisson sampling. For instance, we could assume periodic sampling subject to skips (deletion of samples) each occurring independently with probability q; the latter model leads to a mean square error expression

$$\epsilon^2 = \frac{2q}{\pi}\gamma^{-1} + o(\gamma^{-1}). \tag{4.17}$$

It is our belief that the γ^{-1} leading term is typical of Wiener-Kolmogorov interpolation of a differentiable signal. Of course, a regular sampling pattern at a high sampling rate leads to error-free recovery, so that this statement is applicable only if some degree of randomness is present in the sampling train.

We have been unable to answer to our satisfaction the nature of the poly-

[7] The exact error formula is $\epsilon^2 = (2/\pi)\gamma^{-1}[1 + (2\gamma^{-1}/\pi)]^{-1}$.

gonal interpolation error when $x(t)$ is once differentiable. If we again introduce (4.2) and (4.3) into the general error expression (3.2) and expand R_x as $R_x(u) = 1 - (u^2/2) E\{[\dot{x}(t)]^2\} + o(u^2)$, we see that all the terms cancel mutually except the last; hence we are left with

$$\epsilon^2 = o(\gamma^{-2}). \tag{4.18}$$

Approximation techniques are also useless in attempting to derive asymptotic error expressions for non-differentiable signals. Fortunately, it has been possible to evaluate the error in certain special cases (e.g., wide-sense Markov signals). Here one is generally able to determine the exact mean square expression, from which the asymptotic form is then obtained.

EXAMPLE 4.4 (wide-sense Markov signal with independent skip sampling—polygonal interpolation). We return to (4.2) and (4.3), which enable us to compute the desired mean square error. In order to deal with the denominators comprised of powers of $L_1 + L_{-1}$, we apply the total probability law

$$E(Z) = \sum E(Z \mid C_n) P(C_n),$$

where $\{C_n\}$ constitutes a partition of the probability space. In this instance, we take C_n as the event that time t falls in a sampling interval of length nT, where T is the nominal distance between successive (unskipped) samples. We then have $C_n = \{L_1 + L_{-1} = nT\}$, and as may readily be verified

$$P(C_n) = nq^{n-1}(1-q)^2 \tag{4.19}$$

under the earlier assumption of independent skips each having probability q. The conditional density of L_1, given C_n, is

$$g_1(u \mid C_n) = \begin{cases} (nT)^{-1} & 0 \leqslant u < nT \\ 0 & \text{otherwise.} \end{cases} \tag{4.20}$$

With this information one computes for later use in the total probability formula

$$E\{[\hat{x}(t)]^2 \mid C_n\} = \tfrac{1}{3}[2 + R_x(nT)], \tag{4.21}$$

which holds generally for any R_x. On specializing (4.21) by assigning $R_x(\tau) = \exp(-a \mid \tau \mid)$, multiplying by $P(C_n)$, and summing one obtains

$$E\{[\hat{x}(t)]^2\} = \frac{1}{3}\left[2 + \frac{(1-q)^2 e^{-aT}}{(1-qe^{-aT})^2}\right] = 1 - \frac{1}{3}\left(\frac{1+q}{1-q}\right)(aT) + o(aT). \tag{4.22}$$

The same technique, applied to the computation of $E[x(t)\,\hat{x}(t)]$, yields the closed form expression

$$E[x(t)\,\hat{x}(t)] = \frac{2(1-q)}{aT}\left\{1 - \frac{1-q}{q(aT)}\left[\log\left(\frac{1-qe^{-aT}}{1-q}\right)\right]\right\}. \quad (4.23)$$

An expansion of the logarithm [whose argument is written

$$1 + \left\{\frac{q}{1-q}(1-e^{-aT})\right\}$$

for this purpose] leads to a high sampling rate approximation to $E[x(t)\,\hat{x}(t)]$ identical to the right side of (4.22). The estimates of $E\{[\hat{x}(t)]^2\}$ and $E[x(t)\hat{x}(t)]$, valid for high sampling rates $\gamma = (1-q)/aT$, are entered into the general mean square error formula $\epsilon^2 = 1 - 2E[x(t)\,\hat{x}(t)] + E\{[\hat{x}(t)]^2\}$, thus giving

$$\epsilon^2 = \left(\frac{1+q}{3}\right)\gamma^{-1} + o(\gamma^{-1}). \quad (4.24)$$

The above error invites comparison with the corresponding error associated with use of only one sampling point [see Eq. (3.20)]. Although the mean square error is reduced by a factor of three by interpolation between two sampling points relative to sample-and-hold extrapolation, its dependence on the sampling rate through γ^{-1} remains unchanged. The benefits of employing a second sampling point are therefore relatively small, and are even further minimized if a two-point (causal) extrapolator is considered in place of the interpolator of the present example.

It is noteworthy that the error for twice differentiable $x(t)$ with a polygonal recovery scheme varies as γ^{-4}, whereas the variation is with γ^{-1} for the nondifferentiable signal just considered. We also showed by (3.18) that the mean square error behaved as $o(\gamma^{-2})$ when the signal is (at least) once differentiable. It therefore appears that signal differentiability is a vital error consideration at high sampling rates. This is again seen from the next example, in which we are also able to make error comparisons with Wiener-Kolmogorov noncausal filtering.

EXAMPLE 4.5 (Poisson sampling of a wide-sense Markov signal—polygonal interpolation). Because L_1 and L_{-1} are statistically independent, each having exponential density with parameter β, one can find the joint probability density of $L_1 + L_{-1}$ and L_{-1} to be

$$h(u, v) = \begin{cases} \beta^2 e^{-\beta u} & 0 \leqslant v \leqslant u \\ 0 & \text{otherwise.} \end{cases} \quad (4.25)$$

Knowledge of this joint density permits us to determine $E[x(t)\,\hat{x}(t)]$ by noting that (4.2) consists of two terms having the same expectation, so that

$$E[x(t)\,\hat{x}(t)] = 2\int_0^\infty\!\!\int \frac{(u-v)\,R_x(v)}{u}\,h(u,v)\,du\,dv. \tag{4.26}$$

We integrate first on v, taking advantage of the exponential form of R_x; the second integral can then be treated as a Laplace transform whose transform variable in β. The final result is

$$E[x(t)\,\hat{x}(t)] = 2\gamma\{1 - \gamma\log(1+\gamma^{-1})\}. \tag{4.27}$$

By proceeding in similar fashion we also obtain

$$E\{[\hat{x}(t)]^2\} = \frac{1}{3}\left\{2 + \frac{1}{(1+\gamma^{-1})^2}\right\}, \tag{4.28}$$

where the nondimensional sampling rate $\gamma = \beta/a$ as before. Although the mean square error is directly available from (4.27) and (4.28) above, the effect of γ on error behavior is not transparent. However, the asymptotic expression resulting from expansion of the logarithm in (4.27) and the denominator in (4.28) takes on the relatively simple form

$$\epsilon^2 = \frac{2}{3}\gamma^{-1} - \frac{8}{15}\gamma^{-3} + o(\gamma^{-3}). \tag{4.29}$$

Here we have deliberately carried the γ^{-3} term, which will prove useful in the next example.

For the nondifferentiable signal $x(t)$ considered above, the error is linear in γ^{-1}, as it was in the preceding example likewise concerned with polygonal interpolation of a nondifferentiable $x(t)$. While the error magnitude is not as responsive to sampling rate as for a twice differentiable signal, the performance (4.29) is nevertheless superior to that achieved by a noncausal Wiener-Kolmogorov interpolator. It has been shown elsewhere [Eq. (2.37) of Leneman (1966b)], that for the latter

$$\epsilon^2 = \sqrt{\frac{1}{2}}\,\gamma^{-1/2} + o(\gamma^{-1/2}), \tag{4.30}$$

which is clearly worse than the mean square error (4.29) applicable to the polygonal interpolator depending on only one point to each side of t.

It is also instructive to observe the relation between sampling rate and mean square error when the signal is exactly once differentiable; as we have seen, this error varies as γ^{-4} for a signal which is twice differentiable, and as γ^{-1} (in Examples 4.4 and 4.5) if the signal cannot be differentiated. Since general results in this direction seem to be beyond our grasp, we must again rely on a convenient example.

EXAMPLE 4.6 (Poisson sampling of once differentiable signal—polygonal interpolation). Although we are unable to treat once differentiable signals of general type, we can use the results of the preceding example to handle any signal of rational spectral density having distinct real poles. The correlation R_x of such a signal can be written

$$R_x(\tau) = \sum^{n} A_k \exp(-a_k \mid \tau \mid), \tag{4.31}$$

whose coefficients must satisfy

$$\sum A_k = 1 \qquad \text{and} \qquad \sum A_k a_k = 0 \tag{4.32}$$

because $E\{[x(t)]^2\} = 1$ and the derivative of R_x at the origin exists and is zero by the assumed differentiability (in quadratic mean) of $x(t)$. Moreover, $\sum A_k a_k^3 \leqslant 0$, with equality iff $x(t)$ is twice differentiable.

Let us now compute the mean square error for an R_x furnished by (4.31). Now the mean square recovery error is linear in R_x, so that the formula (4.29) can be applied term-by-term to (4.31). For notational brevity, call the right side of (4.29) $\epsilon^2(\gamma)$. Take $\gamma = \beta/a$, where β is the Poisson rate as before, and a is any consistent measure of signal bandwidth (e.g., the half power frequency). The linearity property, as applied to the mean square error, then yields

$$\epsilon^2 = \sum A_k \epsilon^2 \left(\frac{\gamma}{b_k}\right), \tag{4.33}$$

in which $b_k = a_k/a$. By virtue of the second equality in (4.32) and the definition of $\epsilon^2(\cdot)$ from (4.29), the γ^{-1} terms sum to zero, and there remains

$$\epsilon^2 = -\left[\frac{8}{15} \sum A_k b_k^3\right] \gamma^{-3} + o(\gamma^{-3}). \tag{4.34}$$

We have thus shown that for this (large) class of precisely once differentiable signals the mean square error varies according to γ^{-3}. The same qualitative behavior may be demonstrated for once differentiable signals whose spectral density has multiple poles, and for the skip sampling considered earlier in different connections.

By way of comparison, we again recall that the error relevant to twice differentiable signals varied with γ^{-4}, while that of nondifferentiable signals behaved like γ^{-1}. Wiener-Kolmogorov noncausal filtering of a nondifferentiable signal yielded the still worse figure $\gamma^{-1/2}$. It is also worth noting that utilization of two interpolation points (in place of one) improves the mean square error behavior from γ^{-2} to γ^{-3} for a once differentiable signal, from γ^{-2} to γ^{-4} for a twice differentiable signal, and leaves the factor γ^{-1} unchanged if the signal is not differentiable at all.

V. General Polynomial Interpolation

The interpolation and extrapolation schemes studied in the preceding two sections represent classical Lagrange polynomial interpolation [Davis (1963), Section 2.5] of degree zero and one, respectively. The computational difficulties experienced there discourages direct mean square error calculations for polynomials of higher degree fitted to larger numbers of sample points. Thus, there seems no way to readily determine the asymptotic mean square error incurred by interpolating an m times differentiable signal by a polynomial of degree $n - 1$ passed through n sample points when $m \leqslant n - 1$ and $n \geqslant 3$. However, we can and do derive an error expression of polynomial interpolation valid for $m \geqslant n$ and n arbitrary. The formula obtained will hold for causal as well as noncausal estimators; in either case, the mean square error varies as γ^{-2n}.

Our starting point is the expression for the Cauchy remainder [Davis (1963), Section 3.1] indicative of the difference between an n times differentiable function f and the $n - 1$ degree polynomial P_{n-1} passed through the points $f(t_k)$, $k = 1, 2,\ldots, n$; we have

$$|f(t) - P_{n-1}(t)| = \frac{|f^n(s)|}{n!} \prod_1^n |t - t_k| \tag{5.1}$$

in which the indicated nth derivative is to be taken at some (unspecified) time s, $t_1 \leqslant s \leqslant t_n$. We assume here that the sampling points $t_1 < t_2 < \cdots < t_n$, without requiring that t falls within the interval $[t_1, t_n]$ marked by the sampling time extremes t_1 and t_n. In fact, $t > t_n$ or $t_1 \leqslant t \leqslant t_n$, according as we consider (causal) extrapolation or (noncausal) interpolation.

The error formula (5.1) applies *ipso facto* to randomly timed sampling of an

n times differentiable stochastic signal $x(t)$. We now have $|t - t_k| = L_{j_k}$, where $j_1, ..., j_n$ represent the combination of indices appropriate to the recovery scheme contemplated; all indices are negative for extrapolation, whereas interpolation leads to some positive and some negative indices. Once the substitution by the L_{j_k} is accomplished, one is tempted to square (5.1) (with f replaced by x), and to take the resulting expectation in order to find the mean square error. But unfortunately, one *cannot* claim that

$$E\left\{[x^n(s)]^2 \prod_1^n [L_{j_k}(t)]^2\right\} = E\{[x^n(0)]^2\}E\left\{\prod_1^n [L_{j_k}(0)]^2\right\}, \tag{5.2}$$

since in general the s on the left side of (5.2) depends not only on $x(t)$, but also on $\{t_n\}$.[8] We must therefore resort to a different technique which avoids the dependence problem, although its validity is limited to asymptotically large sampling rates γ.

It is possible to bound the error (5.1) by replacing $f^n(s)$ with

$$\sup_{t_1 \leqslant s \leqslant t_n} [f^n(s)].$$

Correspondingly, we may write in terms of the nondimensional processes introduced in Section II

$$\gamma^{2n}\epsilon^2 \leqslant (n!)^{-2} E\left\{\sup_{*}[\tilde{x}^n(u)]^2 \prod_1^n [\tilde{L}_{j_k}(\tilde{t})]^2\right\}, \tag{5.3}$$

where $*$ indicates that the supremum is to be taken over the interval $[\gamma^{-1}\tilde{t}_1, \gamma^{-1}\tilde{t}_n]$. The right side of (5.3) is further enlarged by taking the supremum instead over the interval $[0 \wedge \gamma^{-1}\tilde{t}_1, 0 \vee \gamma^{-1}\tilde{t}_n]$, in which the symbols $\wedge$ and $\vee$ are defined by $a \wedge b = \min(a, b)$ and $a \vee b = \max(a, b)$, respectively. The supremum of $[\tilde{x}^n(u)]^2$ over the larger interval is denoted by $**$. We now assume the following: the nth derivative of $x(t)$ not only exists, but is also almost surely continuous at each t.[9] A consequence of this assumption is

$$\lim_{\gamma \to \infty}\{\sup_{**}[\tilde{x}^n(u)]^2\} = [\tilde{x}^n(0)]^2; \tag{5.4}$$

here the interval over which the supremum is taken shrinks to the single point zero, so that the indicated convergence is monotone. Lebesgue's monotone

[8] It is not even obvious—although it can be shown—that s is measurable on the product space generated by the stochastic processes $x(t)$ and $\{t_n\}$.

[9] For $n = 1$ and $n = 2$, the results to be derived in this section are already given by (3.11) and (4.5), respectively, but without any requirement on the continuity of the nth derivative.

convergence theorem, applied to (5.3) with $\sup_*$ replaced by $\sup_{**}$, therefore yields

$$\overline{\lim_{\gamma\to\infty}}\, \gamma^{2n}\epsilon^2 \leqslant (n!)^{-2}E\{[\tilde{x}^n(0)]^2\}E\left\{\prod_1^n [\tilde{L}_{j_k}(0)]^2\right\}. \tag{5.5}$$

In writing (5.5), we have taken advantage of the independence of $x(0)$ and $\{t_n\}$, and have used the stationarity of $\{t_n\}$ to eliminate $\tilde{t}$ [Beutler and Leneman (1966a), Theorem 2.3.1).]

The arguments just concluded may be modified by substituting $\inf_*$ for $\sup_*$, in which case the inequalities obtained are in the opposite direction. One may again employ the monotone convergence theorem, which now yields

$$\underline{\lim}_{\gamma\to\infty}\, \gamma^{2n}\epsilon^2 \geqslant (n!)^{-2}E\{[\tilde{x}^n(0)]^2\}E\left\{\prod_1^n [\tilde{L}_{j_k}(0)]^2\right\}. \tag{5.6}$$

It follows from (5.5) and (5.6) above that the limit of $\gamma^{2n}\epsilon^2$ exists and is given by the right side of these inequalities. A statement equivalent to this assertion is

$$\epsilon^2 = \gamma^{-2n}(n!)^{-2}E\{[\tilde{x}^n(0)]^2\}E\left\{\prod_1^n [\tilde{L}_{j_k}(0)]^2\right\} + o(\gamma^{-2n}). \tag{5.7}$$

We have thus shown that the mean square error for polynomial interpolation of a randomly sampled signal is consistent with the classical error formula as well as the results obtained earlier in the context of sample-and-hold and polygonal interpolation recovery.

Summary and Conclusions

We have analyzed the asymptotic mean square errors for polynomial recovery schemes when stochastic signals are sampled at random times. The results are expressed in dimensionless parametric form emphasizing the role of the nondimensional sampling rate γ.

Sample-and-hold recovery is seen to exhibit mean square error proportional to γ^{-2} when the signal is differentiable, and γ^{-1} when it is not. Error penalties for jitter appear to be negligible, nor does low probability random skipping exact a major error increase. Introduction of exponential decay fails to improve high sampling rate performance measurably when the signal is not differentiable; for a differentiable signal, the decay changes the dependence on sampling rate from γ^{-2} to γ^{-1}.

Polygonal interpolation or extrapolation yields a mean square error depending on γ^{-4} for twice differentiable signals. The error varies according to γ^{-3} when the signal is only once differentiable, and with γ^{-1} for a non-differentiable signal. It is clear from these results that signal differentiability is a desirable attribute at high sampling rates.

It is shown that n point polynomial interpolation gives mean square errors proportionate to γ^{-2n} if the signal is at least n times continuously differentiable. The mean square error also varies directly with moments of the forward and/or backward recurrence times, as it does with sample-and-hold and polygonal interpolation; this variation is multiplicative and does not negate the γ dependencies mentioned above.

Virtues of the simple (e.g. sample-and-hold and polygonal) interpolators (or extrapolators) include ease of mechanization, lack of complex design computations, robustness and good performance. Comparisons regarding the latter have been made with Wiener-Kolmogorov filters for the few cases when the optimal time-invariant interpolator (extrapolator) could be derived. Mean square errors for the Wiener-Kolmogorov interpolators tended to vary with $\gamma^{-1/2}$, which distinctly favors the simple interpolators (extrapolators). An explanation of this apparent paradox is easily found; although the simple interpolators are easy to implement, they nevertheless constitute time-varying filters, whereas the Wiener-Kolmogorov filter is constrained to be time-invariant.

RECEIVED: June 28, 1973; REVISED: May 21, 1974

REFERENCES

BALAKRISHNAN, A. V. (1962), On the problem of jitter in sampling, *IRE Trans. Inform. Theory* **8**, 226.

BEUTLER, F. J. (1961), Prediction for wide-sense Markov processes, *IRE Trans. Inform. Theory* **7**, 267.

BEUTLER, F. J. (1963), Multivariate wide-sense Markov processes and prediction theory, *Ann. Math. Statist.* **34**, 424.

BEUTLER, F. J. (1966), Error-free recovery of signals from irregularly spaced samples, *SIAM (Soc. Ind. Appl. Math.) Rev.* **8**, 328.

BEUTLER, F. J. AND LENEMAN, O. (1966a), The theory of stationary point processes, *Acta Math.* **116**, 159.

BEUTLER, F. J. AND LENEMAN, O. (1966b), Random sampling of random processes: stationary point processes, *Inform. Contr.* **9**, 325.

BEUTLER, F. J. AND LENEMAN, O. (1968), The spectral analysis of impulse processes, *Inform. Contr.* **12**, 236.

BEUTLER, F. J. AND LENEMAN, O. (1971), On the statistics of random pulse processes, *Inform. Contr.* **18**, 326.

BROWN, W. M. (1961), Sampling with random jitter, *J. SIAM* **11**, 460.

DAVIS, P. J. (1963), "Interpolation and Approximation," Blaisdell, New York.

GELFAND, I. AND VILENKIN, N. (1964), Applications of Harmonic Analysis, *in* "Generalized Functions" (A. Feinstein, Transl.), Vol. 4, Academic Press, New York.

KRYUKOV, V. I. (1967), Calculation of the correlation function and the spectral power density of random sampling, *Radio Eng. Electronic Phys. (USSR)* **12**.

LENEMAN, O. A. (1966a), On error bounds for jittered sampling, *IEEE Trans. Automat. Contr.* **11**, 150.

LENEMAN, O. A. (1966b), Random sampling of random processes: optimum linear interpolation, *J. Franklin Inst.* **281**, 302.

LENEMAN, O. AND LEWIS, J. (1965), A note on reconstruction for randomly sampled data, *IEEE Trans. Automat. Contr.* **10**, 626.

LENEMAN, O. AND LEWIS, J. (1966a), On mean square reconstruction error, *IEEE Trans. Automat. Contr.* **11**, 324.

LENEMAN, O. AND LEWIS, J. (1966b), Random sampling of random processes: mean-square comparison of various interpolators, *IEEE Trans. Automat. Contr.* **11**, 396.

LIFF, A. I. (1965), Mean square reconstruction error, *IEEE Trans. Automat. Contr.* **10**, 370.

LIU, B. AND STANLEY, T. (1965), Error bounds for jittered sampling, *IEEE Trans. Automat. Contr.* **10**, 449.

LOEVE, M. (1955), "Probability Theory," Van Nostrand, New York.

MCFADDEN, J. A. (1965), The entropy of a point process, *J. SIAM* **13**, 988.

ROZANOV, Y. A. (1959), To the extrapolation of generalized random stationary processes, *Theor. Probability Appl.* **4**, 426.

SHANNON, C. E. (1949), Communications in the presence of noise, *Proc. IRE* **37**, 10.